DOI:
10.1142/S0219493715500069

Document Version
Early version, also known as pre-print

Link to publication record in King's Research Portal

Citation for published version (APA):
Casserini, M., & Liang, G. (2015). Fully coupled forward-backward stochastic dynamics and functional differential systems. *Stochastics and Dynamics*, *15*(2). https://doi.org/10.1142/S0219493715500069

Fully coupled forward-backward stochastic dynamics and functional differential systems

Matteo Casserini* and Gechun Liang†

*Department of Mathematics, ETH Zürich, Switzerland,
matteo.casserini@math.ethz.ch

†Department of Mathematics, King's College London, U.K.,
gechun.liang@kcl.ac.uk

†Oxford-Man Institute, University of Oxford, U.K.

Abstract

This article introduces and solves a general class of fully coupled forward-backward stochastic dynamics by investigating the associated system of functional differential equations. As a consequence, we are able to solve many different types of forward-backward stochastic differential equations (FBSDEs) that do not fit in the classical setting. In our approach, the equations are running in the same time direction rather than in a forward and backward way, and the conflicting nature of the structure of FBSDEs is therefore avoided.

Keywords. Backward stochastic differential equation, BSDE, fully coupled forward-backward stochastic differential equation, FBSDE, functional differential equation, functional differential system.

Mathematics Subject Classification (2010). 60H10, 60H30, 93E03.

This work is partially supported by ETH Zürich and the Oxford-Man Institute (University of Oxford). The authors would like to thank their respective advisors Freddy Delbaen (ETH), Terry Lyons (Oxford) and Zhongmin Qian (Oxford) for their helpful discussions and comments.

1 Introduction

Due to their central role at the intersection between stochastic analysis, mathematical finance and partial differential equations, backward stochastic differential equations (BSDEs) and forward-backward stochastic differential equations (FBSDEs) have been subject of extensive research during the last two decades. While linear BSDEs had already been introduced by Bismut [2] in 1973, it was only after the seminal work of Pardoux and Peng [19] in 1990, who first studied the general non-linear case, that BSDEs gained considerable attention. Since then, the importance of the theory of BSDEs increased dramatically, finding numerous applications in stochastic control theory, PDE theory, mathematical finance and many other fields. We refer the reader to the books [10, 18] and the surveys [9, 11] for an extensive overview of BSDEs and their applications.

While simple types of decoupled FBSDEs had already been considered by Pardoux and Peng [20], the study of fully coupled FBSDEs has been initiated by Antonelli [1], who proved the existence and uniqueness of local solutions. The solvability of fully coupled FBSDEs was later studied by several authors, and mainly three types of methods have been proposed so far, each having its constraints and which do not cover each other. The first is the method of contraction mapping, introduced in the local case by Antonelli [1]: his work was later developed by Pardoux and Tang [21] to solve FBSDEs globally under additional monotonicity conditions. Later on, motivated by the method of continuation in PDE theory, Hu and Peng [13], Peng and Wu [22] and Yong [24] solved FBSDEs on arbitrary intervals by relying on a different type of monotonicity assumptions on the coefficients. Finally, Ma et al. [16] introduced the well-known four-step scheme, which links FBSDEs and quasilinear PDEs: in this case, the coefficients have to be deterministic and satisfy strong regularity assumptions. This method was further developed by Delarue [4], who relaxed the regularity assumption on the coefficients by combining the four-step scheme with the contraction method. In the last years, Zhang [25,26] and more recently Ma et al. [17] developed a so-called decoupling scheme to solve FBSDEs with random coefficients, which unifies most of the existing results at least in the one-dimensional case. For an extensive account of FBSDEs, we refer to the book by Ma and Yong [18].

The purpose of this article is to introduce a general class of fully coupled forward-backward stochastic dynamics on a general filtered probability space, which contains classical FBSDEs as a special case. Inspired by the recent work on Lipschitz BSDEs of Liang et al. [15], we study the solvability of these forward-backward dynamics by introducing an appropriate *functional differential system* of the form

$$\begin{cases} dX_t = \mu(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^1(\mathcal{M}(X,V))_t)dt + \sigma(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^2(\mathcal{M}(X,V))_t)dW_t, \\ dV_t = f(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^3(\mathcal{M}(X,V))_t)dt, \\ X_0 = x, \quad V_0 = 0, \end{cases}$$

where μ, σ, f are random functions, $\mathcal{L}^1$, $\mathcal{L}^2$, $\mathcal{L}^3$ are general abstract operators, and $\mathcal{M}$, $\mathcal{Y}$ are given, for a random function ϕ, by

$$\begin{aligned} \mathcal{M}(X,V)_t &= \mathcal{M}^\phi(X,V)_t := E[\phi(X_T) + V_T|\mathcal{F}_t], \\ \mathcal{Y}(X,V)_t &= \mathcal{Y}^\phi(X,V)_t := \mathcal{M}^\phi(X,V)_t - V_t. \end{aligned}$$

In particular, our approach does not rely a priori on the existence of martingale representations, and shows that FBSDEs can be reformulated as functional differential equations defined in a forward way: since both equations are running in the same time direction, this avoids the conflicting nature of the structure of FBSDEs. More important, our results allow us to consider a more general class of forward-backward stochastic dynamics which are beyond the existing framework, and extend the results of Liang et al. [15] from backward systems to fully coupled forward-backward systems.

After introducing an appropriate framework and defining properly the problem, we study its local solvability and derive our main result: the existence of a unique local solution to the functional differential system under Lipschitz and monotonicity conditions on the coefficients and under specific Lipschitz assumptions on the operators $\mathcal{L}^i$. In particular, the conditions on $\mathcal{L}^i$ are rather mild and allow to consider many types of operators different from the usual martingale integrand processes treated in classical FBSDEs: as a consequence, we can solve within our framework many different types of forward-backward equations that do not fit in the classical FBSDE setting. To emphasize the generality of these assumptions, we present several examples of possible operators and potential financial applications.

In the second part of the article, we discuss the solvability of the system on arbitrarily large time intervals. This is however more problematic: indeed, it appears impossible to study such an extension without defining the operators $\mathcal{L}^i$ explicitly, and one has to consider the problem separately for each choice of $\mathcal{L}^i$. We conclude the article by presenting a study of the case where the filtration is Brownian and the operators $\mathcal{L}^i$ are given by Itô's representation.

The article is organized as follows. First of all, to provide some intuition, we give in Section 2 a brief overview of the functional differential approach in a simple Brownian setting. In Section 3, after introducing a more general framework, we give a rigorous definition of our class of forward-backward dynamics and the associated functional differential system. Section 4 is then dedicated to the existence and uniqueness of solutions to the latter system for sufficiently small time horizons. Finally, in Section 5 we discuss the general problem of extending the solution to arbitrarily large time intervals, and study in particular the case of classical Brownian FBSDEs.

2 The functional differential approach

We would like to begin by providing the reader with some intuition of the approach we are going to use in the sequel, which is inspired by the work of Liang et al. [15]. We first present the following elementary, but very illustrative result derived in [15]:

Remark 2.1 (Liang et al. [15]). For $T > 0$, assume that we have a special semimartingale $(Y_t)_{t\in[0,T]}$ on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\in[0,T]}, P)$ satisfying the usual assumptions, and let the terminal value $Y_T = \xi \in L^1(\mathcal{F}_T)$ be given. Furthermore, assume that the canonical decomposition of Y is given by

$$Y_t = M_t - V_t,$$

where M is a martingale and V a predictable process of finite variation with

$V_0 = 0$. Then, if V_T is integrable, it is easy to verify that, for all $t \in [0,T]$,

$$M_t = E[M_T|\mathcal{F}_t] = E[\xi + V_T|\mathcal{F}_t],$$
$$Y_t = M_t - V_t = E[\xi + V_T|\mathcal{F}_t] - V_t.$$

In other words, the semimartingale Y and the martingale M can be expressed as operators of the terminal value ξ and the finite variation process V.

We show now, with the help of some intuitive arguments, how this remark can lead us to an alternative formulation of the classical FBSDE problem. Let $(\Omega, \mathcal{F}, P)$ be for the moment a complete probability space with an m-dimensional Brownian motion $W = (W_t)_{t\in[0,T]}$ and the corresponding filtration $(\mathcal{F}_t)_{t\in[0,T]}$, augmented by the P-null sets in $\mathcal{F}$. We consider a classical fully coupled FBSDE of the form

$$\begin{cases} dX_t = \mu(t, X_t, Y_t, Z_t)dt + \sigma(t, X_t, Y_t, Z_t)dW_t, \\ dY_t = -f(t, X_t, Y_t, Z_t)dt + Z_t dW_t, \\ X_0 = x, \quad Y_T = \phi(X_T), \end{cases} \tag{2.1}$$

where the functions $\mu : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^n$, $\sigma : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^{n\times m}$, $f : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^d$, $\phi : \Omega \times \mathbb{R}^n \to \mathbb{R}^d$ satisfy the usual measurability and integrability conditions.

Assume now that the above FBSDE has a solution (X, Y, Z). Since $\Phi(X_T)$ is the terminal value of the semimartingale Y, Remark 2.1 induces us to introduce the operators $\mathcal{M}^\phi$ and $\mathcal{Y}^\phi$ by defining

$$\mathcal{M}^\phi(X,V)_t := E[\phi(X_T) + V_T|\mathcal{F}_t],$$
$$\mathcal{Y}^\phi(X,V)_t := \mathcal{M}^\phi(X,V)_t - V_t, \quad t \in [0,T],$$

for any processes X, V such that $\phi(X_T) \in L^1(\mathcal{F}_T)$, $V_T \in L^1(\mathcal{F}_T)$. Then, by Remark 2.1 and the definition of the operators $\mathcal{Y}^\phi$ and $\mathcal{M}^\phi$, it seems plausible to associate the above FBSDE to the following system of *forward functional differential equations*

$$\begin{cases} dX_t = \mu\big(t, X_t, \mathcal{Y}^\phi(X,V)_t, \mathcal{Z}^\phi(X,V)_t\big)dt + \sigma(t, X_t, \mathcal{Y}^\phi(X,V)_t, \mathcal{Z}^\phi(X,V)_t)dW_t, \\ dV_t = f\big(t, X_t, \mathcal{Y}^\phi(X,V)_t, \mathcal{Z}^\phi(X,V)_t\big)dt, \\ X_0 = x, \quad V_0 = 0, \end{cases}$$

where $\mathcal{Z}^\phi$ is given implicitly via Itô's representation theorem by

$$\mathcal{M}^\phi(X,V)_T = E[\mathcal{M}^\phi(X,V)_T] + \int_0^T \mathcal{Z}^\phi(X,V)_s dW_s.$$

The peculiarity of these stochastic differential equations consists in the fact that the coefficients μ, σ, f depend not only on the behaviour of the solution process (X, V) up to the present value, but also on the terminal value (X_T, V_T) of the solution: such stochastic differential equations are not standard, and for this reason they are called *functional* differential equations (note that the term "functional differential equations" is often used in the literature to refer to stochastic delay differential equations: however, contrary to the latter, the drivers of the equations studied here do not have any delay in the past, but rather in the future).

3 Fully coupled functional differential systems

In this section, we will show how the approach presented above can be made rigorous and extended to a much more general framework. To this end, we first need to introduce some notation: in the following, we fix $T > 0$, and assume that we are given a complete probability space $(\Omega, \mathcal{F}, P)$ together with a general filtration $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ satisfying the usual assumptions. We recall that every $\mathbb{F}$-martingale has under these conditions a càdlàg version, which we will always choose. Moreover, we denote by $\mathcal{P}$ the predictable σ-field with respect to $\mathbb{F}$ and assume that an m-dimensional Brownian motion $W = (W_t)_{t\in[0,T]}$ is defined on $(\Omega, \mathcal{F}, \mathbb{F}, P)$.

For k, $l \in \mathbb{N}$, $|\cdot|$ denotes the Euclidean norm on $\mathbb{R}^k$, respectively the Hilbert-Schmidt norm on $\mathbb{R}^{k\times l}$, and $\mathbb{R}^{k\times l}$ will often be identified with $\mathbb{R}^{k\cdot l}$. We define $\mathscr{S}^2([0,T], \mathbb{R}^d)$ as the space of all processes $V : \Omega \times [0,T] \to \mathbb{R}^d$ continuous and adapted such that $V_0 = 0$ and $E[\sup_{t\in[0,T]} |V_t|^2] < \infty$, while $\mathscr{M}^2([0,T], \mathbb{R}^d)$ denotes the space of all square integrable $\mathbb{R}^d$-valued martingales on $[0,T]$. Both $\mathscr{S}^2([0,T], \mathbb{R}^d)$ and $\mathscr{M}^2([0,T], \mathbb{R}^d)$ are endowed with the norm

$$\|V\|_{\mathscr{S}^2[0,T]} := \sqrt{E\Big[\sup_{t\in[0,T]} |V_t|^2\Big]},$$

and note that $\big(\mathscr{S}^2([0,T], \mathbb{R}^d), \|\cdot\|_{\mathscr{S}^2[0,T]}\big)$ is then a Banach space. Sometimes, we will also need the direct sum space $\mathscr{S}^2([0,T], \mathbb{R}^d) \oplus \mathscr{M}^2([0,T], \mathbb{R}^d)$, endowed with the same norm $\|\cdot\|_{\mathscr{S}^2[0,T]}$.

In the following, we will consider a particular generalization of the FBSDE (2.1) on the filtered probability space $(\Omega, \mathcal{F}, \mathbb{F}, P)$. More exactly, we will assume that μ, σ and f depend on some general processes $\mathcal{L}^i(M)$, $i = 1, 2, 3$, instead than on Z, where $\mathcal{L}^i$ is for each i an abstract operator defined on $\mathscr{M}^2([0,T], \mathbb{R}^d)$ and taking values, for some $p_i \in \mathbb{N}$, in the space of p_i-dimensional adapted processes (the codomains of $\mathcal{L}^i$ will be further specified later). As we will see later more in detail, this substitution will allow both to take into account the generality of the filtration $\mathbb{F}$ and to treat locally other types of forward-backward equations not fitting in the classical framework. This generalization of the FBSDE (2.1) leads us to the following system:

$$\begin{cases} dX_t = \mu(t, X_t, Y_t, \mathcal{L}^1(M)_t)dt + \sigma(t, X_t, Y_t, \mathcal{L}^2(M)_t)dW_t, \\ dY_t = -f(t, X_t, Y_t, \mathcal{L}^3(M)_t)dt + dM_t, \\ X_0 = x, \quad Y_T = \phi(X_T), \end{cases} \tag{3.1}$$

where $\mu : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{p_1} \to \mathbb{R}^n$, $\sigma : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{p_2} \to \mathbb{R}^{n\times m}$, $f : \Omega \times [0,T] \times \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{p_3} \to \mathbb{R}^d$, $\phi : \Omega \times \mathbb{R}^n \to \mathbb{R}^d$ have the necessary measurability and integrability properties. Note that the system is then completely determined by $(\mu, \sigma, f, \phi, \mathcal{L}^1, \mathcal{L}^2, \mathcal{L}^3)$.

Definition 3.1. *A solution to* (3.1) *is a triplet of processes* (X, Y, M) *such that* $X \in \mathscr{S}^2([0,T], \mathbb{R}^n)$, $Y \in \mathscr{S}^2([0,T], \mathbb{R}^d) \oplus \mathscr{M}^2([0,T], \mathbb{R}^d)$, $M \in \mathscr{M}^2([0,T], \mathbb{R}^d)$, *and satisfying the integral formulation of* (3.1)*.*

We will call such systems *fully coupled forward-backward stochastic dynamics*. As mentioned in the previous section, a viable approach to study the solvability of this system is to reformulate the problem with the help of appropriate functional differential equations. To this end, we denote by $\mathscr{C}_X^\phi$ the class of $\mathbb{R}^n$-valued adapted processes X on $[0,T]$ such that $\phi(X_T) \in L^1(\mathcal{F}_T)$, by $\mathscr{C}_V$ the class of $\mathbb{R}^d$-valued adapted processes V on $[0,T]$ such that $V_T \in L^1(\mathcal{F}_T)$, and we define the operators $\mathcal{M}^\phi$ and $\mathcal{Y}^\phi$ on $\mathscr{C}_X^\phi \times \mathscr{C}_V$ by

$$\begin{aligned} \mathcal{M}^\phi(X,V)_t &:= E[\phi(X_T) + V_T|\mathcal{F}_t], \\ \mathcal{Y}^\phi(X,V)_t &:= \mathcal{M}^\phi(X,V)_t - V_t, \quad t \in [0,T]. \end{aligned} \tag{3.2}$$

For the rest of this article, we will drop the dependence of $\mathcal{M}^\phi$ and $\mathcal{Y}^\phi$ on ϕ by writing $\mathcal{M}$ and $\mathcal{Y}$. With the help of the operators $\mathcal{M}$ and $\mathcal{Y}$, we reformulate the problem (3.1) as the following fully coupled system of *forward functional differential equations*:

$$\begin{cases} dX_t = \mu(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^1(\mathcal{M}(X,V))_t)dt + \sigma(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^2(\mathcal{M}(X,V))_t)dW_t, \\ dV_t = f(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^3(\mathcal{M}(X,V))_t)dt, \\ X_0 = x, \quad V_0 = 0. \end{cases} \tag{3.3}$$

Definition 3.2. *A solution to* (3.3) *is a pair of processes* (X,V) *such that* $(X,V) \in \mathscr{S}^2([0,T],\mathbb{R}^n) \times \mathscr{S}^2([0,T],\mathbb{R}^d)$ *and satisfying the integral formulation of* (3.3).

Such systems will be called *fully coupled functional differential systems.* It is not difficult to show the equivalence of the systems (3.1) and (3.3):

Lemma 3.3. *The fully coupled forward-backward system* (3.1) *has a solution if and only if the functional differential system* (3.3) *does.*

Proof. If (X,Y,M) solves (3.1), then we obtain a solution of (3.3) via the canonical decomposition of the semimartingale Y. Conversely, if (X,V) is a solution of the functional differential system, then $\big(X, \mathcal{Y}(X,V), \mathcal{M}(X,V)\big)$ solves (3.1). □

An immediate observation is that the functional differential system (3.3) has a more homogeneous structure than that of the original problem. Indeed, while the forward-backward dynamics (3.1) consist of a forward and a backward equation of different nature, both the functional differential equations in (3.3) are running forward in time and show a similar dependence of the coefficients on both the present and the terminal values of the solution processes. In particular, this homogeneity allows to rewrite the problem more compactly as

$$\begin{cases} d\mathcal{U}_t = \Psi(t, \pi_1(\mathcal{U}_t), \mathcal{Y}(\mathcal{U})_t, \mathcal{L}^1(\mathcal{M}(\mathcal{U}))_t, \mathcal{L}^3(\mathcal{M}(\mathcal{U}))_t)dt \\ \qquad\qquad + \Sigma(t, \pi_1(\mathcal{U}_t), \mathcal{Y}(\mathcal{U})_t, \mathcal{L}^2(\mathcal{M}(\mathcal{U}))_t)dW_t, \\ \mathcal{U}_0 = (x,0)^{\mathrm{T}}, \end{cases}$$

where $\mathcal{U} = (X,V)^{\mathrm{T}}$, $\pi_1 : \mathbb{R}^{n+d} \to \mathbb{R}^n$ is the projection on the first n components, $\Sigma = (\sigma, 0)^{\mathrm{T}}$, and for $t \in [0,T]$, $(x,y,z_1,z_3) \in \mathbb{R}^n \times \mathbb{R}^d \times \mathbb{R}^{p_1} \times \mathbb{R}^{p_3}$, $\Psi(t,x,y,z_1,z_3) = (\mu(t,x,y,z_1), f(t,x,y,z_3))^{\mathrm{T}}$. For the rest of the article, we prefer however to consider the system as formulated in (3.3), since it will be more convenient to treat the coupling between X and V in such a framework.

4 Existence and uniqueness of local solutions

Because of Lemma 3.3, we will now focus our attention on the existence and uniqueness of solutions to the fully coupled functional differential system (3.3). As a first step, we study the local solvability of the problem and introduce sufficient monotonicity and Lipschitz assumptions on the coefficients μ, σ, f, ϕ and the operators $\mathcal{L}^1$, $\mathcal{L}^2$ and $\mathcal{L}^3$.

In the following, we denote by $\mathscr{H}^2([0,T],\mathbb{R}^l)$ the space of $\mathcal{P}$-measurable processes $H:\Omega\times[0,T]\to\mathbb{R}^l$ such that $\|H\|^2_{\mathscr{H}^2[0,T]} := E[\int_0^T |H_t|^2 dt] < \infty$, where $l\in\mathbb{N}$. First of all, we shall assume that the coefficients μ, σ, f, ϕ satisfy the following assumption:

Assumption (A1): *The functions $\mu:\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_1}\to\mathbb{R}^n$, $\sigma:\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_2}\to\mathbb{R}^{n\times m}$, $f:\Omega\times[0,T]\times\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_3}\to\mathbb{R}^d$ and $\phi:\Omega\times\mathbb{R}^n\to\mathbb{R}^d$ satisfy Assumption* **(A1)** *if there exists a constant $C>0$ such that:*

(A1.1) *For any $(x,y,z_1,z_2,z_3)\in\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_1}\times\mathbb{R}^{p_2}\times\mathbb{R}^{p_3}$, the processes $\mu(\cdot,x,y,z_1)$, $\sigma(\cdot,x,y,z_2)$ and $f(\cdot,x,y,z_3)$ are $\mathcal{P}$-measurable and $\phi(x)$ is $\mathcal{F}_T$-measurable.*

(A1.2) *For every $(x,y,z_1),(x',y',z_1')\in\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_1}$,*

$$\begin{aligned}(x-x')^{\mathrm{T}}\big(\mu(\cdot,x,y,z_1)-\mu(\cdot,x',y,z_1)\big) &\le C|x-x'|^2,\\ |\mu(\cdot,x,y,z_1)-\mu(\cdot,x,y',z_1')| &\le C(|y-y'|+|z_1-z_1'|),\\ |\mu(\cdot,x,0,0)| &\le C(1+|x|) \quad dP\otimes dt\text{-a.s.},\end{aligned}$$

and the function $x\mapsto\mu(\cdot,x,y,z_1)$ is $dP\otimes dt$-a.s. continuous.

(A1.3) *$f(\cdot,0,0,0)\in\mathscr{H}^2([0,T],\mathbb{R}^d)$, $\sigma(\cdot,0,0,0)\in\mathscr{H}^2([0,T],\mathbb{R}^{n\times m})$ and $\phi(0)\in L^2(\Omega,\mathbb{R}^d)$.*

(A1.4) *For every $(x,y,z_2,z_3),(x',y',z_2',z_3')\in\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_2}\times\mathbb{R}^{p_3}$,*

$$\begin{aligned}|\sigma(\cdot,x,y,z_2)-\sigma(\cdot,x',y',z_2')|^2 &\le C(|x-x'|^2+|y-y'|^2+|z_2-z_2'|^2),\\ |f(\cdot,x,y,z_3)-f(\cdot,x',y',z_3')| &\le C(|x-x'|+|y-y'|+|z_3-z_3'|),\\ |\phi(x)-\phi(x')| &\le C|x-x'| \quad dP\otimes dt\text{-a.s.}.\end{aligned}$$

Remark 4.1. The above conditions on μ, σ, f and ϕ are quite standard in the theory of FBSDEs (see for instance [18]). Moreover, the reader can easily verify that the condition (A1.2) could be replaced by the following stronger, but more standard assumption:

(A1.2') $\mu(\cdot,0,0,0)\in\mathscr{H}^2([0,T],\mathbb{R}^n)$ *and, for $(x,y,z_1),(x',y',z_1')\in\mathbb{R}^n\times\mathbb{R}^d\times\mathbb{R}^{p_1}$,*

$$|\mu(t,x,y,z_1)-\mu(t,x',y',z_1')|\le C(|x-x'|+|y-y'|+|z_1-z_1'|)\ \ dP\otimes dt\text{-a.s.}.$$

In particular, the assumptions on ϕ allow us to derive the following Lipschitz estimates on the operators $\mathcal{Y}$ and $\mathcal{M}$, which will be essential in the sequel:

Lemma 4.2. *Assume that ϕ satisfies the conditions in* **(A1)**. *Then we have that, for the operators introduced in* (3.2),

$$\begin{aligned}\mathcal{Y} &: \mathscr{S}^2([0,T],\mathbb{R}^n)\times\mathscr{S}^2([0,T],\mathbb{R}^d)\to\mathscr{S}^2([0,T],\mathbb{R}^d)\oplus\mathscr{M}^2([0,T],\mathbb{R}^d),\\ \mathcal{M} &: \mathscr{S}^2([0,T],\mathbb{R}^n)\times\mathscr{S}^2([0,T],\mathbb{R}^d)\to\mathscr{M}^2([0,T],\mathbb{R}^d).\end{aligned}$$

Moreover, for any $X, X'\in\mathscr{S}^2([0,T],\mathbb{R}^n)$ and $V,V'\in\mathscr{S}^2([0,T],\mathbb{R}^d)$,

$$\begin{aligned}\|\mathcal{Y}(X,V)-\mathcal{Y}(X',V')\|_{\mathscr{S}^2[0,T]} &\le 2C\|X-X'\|_{\mathscr{S}^2[0,T]}+3\|V-V'\|_{\mathscr{S}^2[0,T]},\\ \|\mathcal{M}(X,V)-\mathcal{M}(X',V')\|_{\mathscr{S}^2[0,T]} &\le 2C\|X-X'\|_{\mathscr{S}^2[0,T]}+2\|V-V'\|_{\mathscr{S}^2[0,T]}.\end{aligned}$$

Proof. We first prove the second assertion. By the triangle inequality,

$$\begin{aligned}&\|\mathcal{M}(X,V)-\mathcal{M}(X',V')\|_{\mathscr{S}^2[0,T]}\\ &\le \big\|E[\phi(X_T)-\phi(X'_T)|\mathcal{F}_\cdot]\big\|_{\mathscr{S}^2[0,T]}+\big\|E[V_T-V'_T|\mathcal{F}_\cdot]\big\|_{\mathscr{S}^2[0,T]}\\ &= E\Big[\sup_{t\in[0,T]}\big|E[\phi(X_T)-\phi(X'_T)|\mathcal{F}_t]\big|^2\Big]^{1/2}+E\Big[\sup_{t\in[0,T]}\big|E[V_T-V'_T|\mathcal{F}_t]\big|^2\Big]^{1/2}\end{aligned}$$

and therefore, by Doob's inequality and the assumption on ϕ,

$$\begin{aligned}\|\mathcal{M}(X,V)-\mathcal{M}(X',V')\|_{\mathscr{S}^2[0,T]} &\le 2\Big(E\big[|\phi(X_T)-\phi(X'_T)|^2\big]^{1/2}+E\big[|V_T-V'_T|^2\big]^{1/2}\Big)\\ &\le 2C\|X-X'\|_{\mathscr{S}^2[0,T]}+2\|V-V'\|_{\mathscr{S}^2[0,T]}.\end{aligned}$$

The estimate for $\|\mathcal{Y}(X,V)-\mathcal{Y}(X',V')\|_{\mathscr{S}^2[0,T]}$ then follows by the application of the triangle inequality. ☐

The next step consists in introducing appropriate conditions for the abstract operators $\mathcal{L}^1$, $\mathcal{L}^2$ and $\mathcal{L}^3$. These are given by the following Lipschitz and boundedness assumptions, which are the same as introduced by Liang et al. [15].

Assumption (L1): *The operator $\mathcal{L}$ satisfies Assumption* **(L1)** *if:*

(L1.1) *$\mathcal{L}$ maps $\mathscr{M}^2([0,T],\mathbb{R}^d)$ into $\mathscr{O}^2([0,T],\mathbb{R}^p)$, where $\mathscr{O}^2([0,T],\mathbb{R}^p)$ is either the space $\mathscr{H}^2([0,T],\mathbb{R}^p)$ or $\mathscr{S}^2([0,T],\mathbb{R}^p)$.*

(L1.2) *$\mathcal{L}$ is bounded and Lipschitz continuous, i.e. there exists a constant $K>0$ independent of T such that, for all $M,M'\in\mathscr{M}^2([0,T],\mathbb{R}^d)$,*

$$\begin{aligned}\|\mathcal{L}(M)\|_{\mathscr{O}^2[0,T]} &\le K\|M\|_{\mathscr{S}^2[0,T]},\\ \|\mathcal{L}(M)-\mathcal{L}(M')\|_{\mathscr{O}^2[0,T]} &\le K\|M-M'\|_{\mathscr{S}^2[0,T]}.\end{aligned}$$

We will see in Theorem 4.5 that Assumption **(L1)** is enough to guarantee the local solvability of our system without a concrete specification of the operators $\mathcal{L}^i$. In particular, as anticipated in the Introduction, the weakness of Assumption **(L1)** allows to consider many different types of operators within our framework: we emphasize its generality by giving several examples of possible operators, and we present potential financial applications.

Examples 4.3. We begin with the classical case of integrand processes generated by martingale representations.

(i) Assume that $(\mathcal{F}_t)_{t\in[0,T]}$ is the augmented filtration generated by the Brownian motion W, and take $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$. Then, we define $\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}^d) \to \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$ implicitly via Itô's representation theorem by

$$M_t = M_0 + \int_0^t \mathcal{L}(M)_s dW_s, \quad t \in [0,T].$$

By Itô's isometry we have that

$$\begin{aligned}\|\mathcal{L}(M)\|^2_{\mathscr{H}^2[0,T]} &= E\left[\int_0^T |\mathcal{L}(M)_t|^2 dt\right] \\ &= E\left[\left(M_T - M_0\right)^2\right] = E[M_T^2] - E[M_0^2] \le \|M\|^2_{\mathscr{S}^2[0,T]},\end{aligned}$$

and the Lipschitz property follows by the linearity of $\mathcal{L}$. Note that in the case where $\mathcal{L}^1 = \mathcal{L}^2 = \mathcal{L}^3 = \mathcal{L}$, the system (3.1) is reduced to a classical FBSDE: this shows that classical FBSDEs can be seen as a special case of the forward-backward stochastic dynamics (3.1).

(ii) Let now $(\mathcal{F}_t)_{t\in[0,T]}$ be a general filtration with just the usual assumptions. As in the previous case, we can take $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$, and define $\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}^d) \to \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$ implicitly via the orthogonal decomposition with respect to W, i.e.

$$M_t = \int_0^t \mathcal{L}(M)_s dW_s + N_t, \quad t \in [0,T],$$

where N is some martingale orthogonal with respect to W. The reader may easily notice the connection between this operator and generalized BSDEs (see [10]), and we can prove similarly to Example 4.3 (i) that $\mathcal{L}$ satisfies **(L1)**, by applying the Burkholder-Davis-Gundy inequality instead of Itô's isometry.

We can thus study generalized fully coupled FBSDEs within our framework, and the choice of the filtration allows us to consider, in typical financial applications, the case of incomplete markets. An illustrative example is that of a large investor trading in an incomplete market: since this investor buys and sells large amounts of assets, it is reasonable to assume that his trading strategy affects the prices of the stocks. By considering the corresponding hedging problem, we thus obtain a fully coupled system, and the incompleteness of the market leads to a generalized fully coupled FBSDE. For more details, we refer the reader to [18].

Examples 4.4. While martingale integrand processes are the case most studied in the literature, they are not the only class of operators fitting in our framework: there are indeed several other classes of non-local operators, not considered in the classical FBSDE literature, which satisfy Assumption **(L1)**. Let us give some examples.

(i) Assume that $(\mathcal{F}_t)_{t\in[0,T]}$ just satisfies the usual assumptions. We take $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{S}^2([0,T],\mathbb{R}^d)$, and $\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}^d) \to \mathscr{S}^2([0,T],\mathbb{R}^d)$ is simply defined by $\mathcal{L}(M) := M$; in this case, Assumption **(L1)** becomes trivial. In a financial context, $\mathcal{L}(M)$ may represent the risky part of the claim $Y = M - V$.

(*ii*) Let for simplicity $d = 1$. Fix $\widetilde{T} > 0$ and assume that, for $T \leq \widetilde{T}$, $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is such that all martingales with respect to $\mathbb{F}$ are continuous. Let $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{H}^2([0,T],\mathbb{R})$, and define $\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}) \to \mathscr{H}^2([0,T],\mathbb{R})$ by

$$\mathcal{L}(M)_t := \sqrt{E\big[\langle M\rangle_{t,T}\big|\mathcal{F}_t\big]}, \quad M \in \mathscr{M}^2([0,T],\mathbb{R}),\ t \in [0,T],$$

where for notational simplicity $\langle M\rangle_{t,T} := \langle M\rangle_T - \langle M\rangle_t$. Then, by the Kunita-Watanabe and the conditional Cauchy-Schwarz inequalities, we have that

$$\begin{aligned} E\big[\langle M,M'\rangle_{t,T}\big|\mathcal{F}_t\big] &\leq E\big[|\langle M,M'\rangle_{t,T}|\big|\mathcal{F}_t\big] \leq E\Big[\sqrt{\langle M\rangle_{t,T}\langle M'\rangle_{t,T}}\Big|\mathcal{F}_t\Big] \\ &\leq \sqrt{E[\langle M\rangle_{t,T}|\mathcal{F}_t]}\sqrt{E[\langle M'\rangle_{t,T}|\mathcal{F}_t]}, \end{aligned}$$

and therefore, by the bilinearity of $\langle\cdot\rangle_{t,T}$,

$$\begin{aligned} \big|\mathcal{L}(M)_t - \mathcal{L}(M')_t\big|^2 &= E[\langle M\rangle_{t,T}|\mathcal{F}_t] + E[\langle M'\rangle_{t,T}|\mathcal{F}_t] \\ &\quad - 2\sqrt{E[\langle M\rangle_{t,T}|\mathcal{F}_t]}\sqrt{E[\langle M'\rangle_{t,T}|\mathcal{F}_t]} \\ &\leq E[\langle M - M'\rangle_{t,T}|\mathcal{F}_t] \end{aligned}$$

for all $t \geq 0$. Hence, by Fubini's theorem,

$$\begin{aligned} \|\mathcal{L}(M) - \mathcal{L}(M')\|^2_{\mathscr{H}^2[0,T]} &= E\Big[\int_0^T \big|\mathcal{L}(M)_t - \mathcal{L}(M')_t\big|^2 dt\Big] \\ &\leq E\Big[\int_0^T E\big[\langle M - M'\rangle_{t,T}\big|\mathcal{F}_t\big]\,dt\Big] \\ &= E\Big[\int_0^T \langle M - M'\rangle_{t,T}\,dt\Big] \leq \widetilde{T}E[\langle M - M'\rangle_T]. \end{aligned}$$

By applying the Burkholder-Davis-Gundy inequality, we finally get the desired Lipschitz property. The boundedness condition is obtained via similar computations.

(*iii*) We choose again for simplicity $d = 1$, and we assume that $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ is as in Example 4.4 (*ii*). We can define $\mathcal{L}$ by taking $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{S}^2([0,T],\mathbb{R})$, and

$$\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}) \to \mathscr{S}^2([0,T],\mathbb{R}), \quad \mathcal{L}(M)_t := \sqrt{\langle M\rangle_t},\ t \in [0,T].$$

By the Burkholder-Davis-Gundy inequality, we have that

$$\|\mathcal{L}(M)\|^2_{\mathscr{S}^2[0,T]} = E\Big[\sup_{t\in[0,T]}\langle M\rangle_t\Big] = E[\langle M\rangle_T] \leq K\|M\|^2_{\mathscr{S}^2[0,T]}.$$

On the other hand, by applying the Kunita-Watanabe inequality,

$$\begin{aligned} \Big|\sqrt{\langle M\rangle_t} - \sqrt{\langle M'\rangle_t}\Big|^2 &= \langle M\rangle_t + \langle M'\rangle_t - 2\sqrt{\langle M\rangle_t\langle M'\rangle_t} \\ &\leq \langle M\rangle_t + \langle M'\rangle_t - 2\big|\langle M,M'\rangle_t\big| \\ &\leq \langle M\rangle_t + \langle M'\rangle_t - 2\langle M,M'\rangle_t = \langle M - M'\rangle_t, \end{aligned}$$

and the Lipschitz property then follows by applying the Burkholder-Davis-Gundy inequality as above. We restrict for a moment to the Brownian setting to give a financial interpretation: in this case, $\mathcal{L}(M)$ can be explicitly rewritten as $\mathcal{L}(M)_t = \sqrt{\int_0^t |Z_s|^2 ds}$, where Z is the martingale integrand in the Itô representation of M. In the usual BSDE framework for hedging (see for instance [11]), $\mathcal{L}(M)$ is then closely connected to the accumulated cost of the portfolio strategy: this could allow us, for instance, to consider storage problems within our setting.

(*iv*) We modify the previous example by combining it with orthogonal decompositions. Let $d = 1$ and $\mathbb{F} = (\mathcal{F}_t)_{t\in[0,T]}$ as in Example 4.4 (*ii*). Fix a martingale $\widetilde{M}$, and define $\mathcal{R} : \mathscr{M}^2([0,T],\mathbb{R}) \to \mathscr{M}^2([0,T],\mathbb{R})$ as the orthogonal term in the orthogonal decomposition with respect to $\widetilde{M}$, i.e.

$$M_t = \big(M_t - \mathcal{R}(M)_t\big) + \mathcal{R}(M)_t, \quad t \in [0,T],$$

where $\mathcal{R}(M)$ is orthogonal with respect to $\widetilde{M}$ and $M_t - \mathcal{R}(M)_t = \int_0^t Z_s d\widetilde{M}_s$ for some process Z. Then, we define the operator $\mathcal{L}$ by

$$\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}) \to \mathscr{S}^2([0,T],\mathbb{R}), \quad \mathcal{L}(M)_t := \sqrt{\langle \mathcal{R}(M)\rangle_t},\ t \in [0,T].$$

Because of the orthogonality, it is easy to check that, for all $t \geq 0$,

$$\begin{aligned}\langle M\rangle_t &= \langle M - \mathcal{R}(M) + \mathcal{R}(M)\rangle_t \\ &= \langle M - \mathcal{R}(M)\rangle_t + \langle \mathcal{R}(M)\rangle_t \geq \langle \mathcal{R}(M)\rangle_t,\end{aligned}$$

and similarly for $M - M'$. Therefore, $\mathcal{L}$ satisfies Assumption **(L1)** because of Example 4.4 (*iii*). This operator may have interesting applications in mathematical finance: namely, in the typical BSDE framework for hedging in incomplete markets, the operator $\mathcal{R}(M)$ represents the non-hedgeable part of the claim. While we cannot hedge this risk, we can partly incorporate its effect on the price process since the coefficients μ, σ and f are allowed to depend on $\sqrt{\langle \mathcal{R}(M)\rangle_t}$.

(*v*) As a final example, we introduce an operator intimately related to backward equations with time delayed generators (see [7]). Let $\widetilde{T} > 0$ be fixed. For $T \leq \widetilde{T}$, let $(\mathcal{F}_t)_{t\in[0,T]}$ satisfy the usual assumptions, and $\mathscr{O}^2([0,T],\mathbb{R}^p) = \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$. Motivated by the framework introduced by Dos Reis et al. [8], we can then define $\mathcal{L} : \mathscr{M}^2([0,T],\mathbb{R}^d) \to \mathscr{H}^2([0,T],\mathbb{R}^{d\times m})$ by

$$\mathcal{L}(M)_t := \int_{-t}^0 \widehat{\mathcal{L}}(M)_{t+s}\alpha_Z(ds),$$

where $\widehat{\mathcal{L}}$ is the operator introduced in Example 4.3 (*ii*), and α_Z is a non-random finite measure with support in $[-\widetilde{T}, 0]$. We observe that $\mathcal{L}$ is closely related to the operator of Example 4.4 (*iii*) when α_Z is the Lebesgue measure restricted to $[-\widetilde{T}, 0]$. Moreover, by applying the change of integration order proved in [8], we can show that

$$\begin{aligned}\|\mathcal{L}(M)\|_{\mathscr{H}^2[0,T]} &= E\left[\int_0^T \left|\int_{-t}^0 \widehat{\mathcal{L}}(M)_{t+s}\alpha_Z(ds)\right|^2 dt\right]^{1/2} \\ &\leq \alpha_Z([-\widetilde{T},0])\|\widehat{\mathcal{L}}(M)\|_{\mathscr{H}^2[0,T]} \leq K\|M\|_{\mathscr{S}^2[0,T]}\end{aligned} \tag{4.1}$$

for some constant $K = K(\widetilde{T})$, where the last inequality follows from the boundedness of $\widehat{\mathcal{L}}$. Assumption **(L1)** then follows by linearity: this will therefore allow us to consider a special class of fully coupled forward-backward equations with delayed generators.

However, the treatment of backward stochastic delayed equations in their full generality would require to replace $\mathcal{Y}(X, V)$ in (3.3) by a new operator $\mathbb{Y}(X, V)_t := \int_{-t}^{0} \mathcal{Y}(X, V)_{t+s}\alpha_Y(ds)$, where α_Y is a non-random finite measure: as pointed out in [8], this kind of equation is very difficult to study even in the simple decoupled case, and it will not be treated here. For an overview of the several applications of backward equations with time delay to mathematical finance, the reader may consult for instance the article by Delong [6].

The broad generality of the above examples should help us understand the importance of the next theorem. This is the main result of this section, and gives the existence of a unique square-integrable solution to the functional differential system (3.3) on sufficiently small intervals $[0, T]$, provided that the operators $\mathcal{L}^i$ satisfy Assumption **(L1)**. This gives us a great flexibility in the choice of $\mathcal{L}^i$, allowing to study many different types of fully coupled, non-classical forward-backward stochastic dynamics.

Theorem 4.5. *Let μ, σ, f and ϕ satisfy Assumption* **(A1)** *with respect to the constant C. Furthermore, assume that $\mathcal{L}^1$, $\mathcal{L}^3$ satisfy Assumption* **(L1)** *and that $\mathcal{L}^2$ satisfies* **(L1)** *with respect to $\mathscr{O}^2([0, T], \mathbb{R}^{p_2}) = \mathscr{S}^2([0, T], \mathbb{R}^{p_2})$, denoting by K the common Lipschitz constant of $\mathcal{L}^1$, $\mathcal{L}^2$, $\mathcal{L}^3$. Then there is a constant $\ell = \ell(C, K)$ depending only on C and K so that, for $T < \ell$,* (3.3) *admits a unique solution (X, V) in $\mathscr{S}^2([0, T], \mathbb{R}^n) \times \mathscr{S}^2([0, T], \mathbb{R}^d)$.*

Proof. From now on, we write $\|\cdot\|_{\mathscr{O}}$ for the norm $\|\cdot\|_{\mathscr{O}^2[0,T]}$ for notational simplicity. Moreover, we denote the product space $\mathscr{S}^2([0, T], \mathbb{R}^n) \times \mathscr{S}^2([0, T], \mathbb{R}^d)$ by $\mathscr{S}^2_X \times \mathscr{S}^2_V$, and endow it with the norm

$$\|(X, V)\|_{\mathscr{S}^2_X \times \mathscr{S}^2_V} := \sqrt{\|X\|^2_{\mathscr{S}^2} + \|V\|^2_{\mathscr{S}^2}}, \quad (X, V) \in \mathscr{S}^2_X \times \mathscr{S}^2_V.$$

The mapping $\mathbb{L} : \mathscr{S}^2_X \times \mathscr{S}^2_V \to \mathscr{S}^2_X \times \mathscr{S}^2_V$, $\mathbb{L}(X, V) := (\widetilde{X}, \widetilde{V})$, is defined as follows: first, $\widetilde{X}$ is constructed as the unique solution in $\mathscr{S}^2([0, T], \mathbb{R}^n)$ to the forward stochastic differential equation

$$\begin{cases} d\widetilde{X}_t = \mu(t, \widetilde{X}_t, \mathcal{Y}(X, V)_t, \mathcal{L}^1(\mathcal{M}(X, V))_t)dt + \sigma(t, \widetilde{X}_t, \mathcal{Y}(X, V)_t, \mathcal{L}^2(\mathcal{M}(X, V))_t)dW_t, \\ \widetilde{X}_0 = x. \end{cases} \tag{4.2}$$

Then, once $\widetilde{X}$ has been obtained, $\widetilde{V}$ is given explicitly by the expression

$$\widetilde{V}_t = \int_0^t f(s, \widetilde{X}_s, \mathcal{Y}(X, V)_s, \mathcal{L}^3(\mathcal{M}(X, V))_s)ds. \tag{4.3}$$

We show that the mapping $\mathbb{L}$ is well defined and maps $\mathscr{S}^2_X \times \mathscr{S}^2_V$ into itself. First of all, we note that the existence of a unique solution in $\mathscr{S}^2([0, T], \mathbb{R}^n)$ to (4.2) follows by Assumption **(A1)** and the results on stochastic differential equations with monotonous coefficients obtained, for instance, by Rozovsky [23]:

indeed, by setting $\bar{\mu}(t,x) := \mu(t,x,\mathcal{Y}(X,V)_t,\mathcal{L}^1(\mathcal{M}(X,V))_t)$ and $\bar{\sigma}(t,x) := \sigma(t,x,\mathcal{Y}(X,V)_t,\mathcal{L}^2(\mathcal{M}(X,V))_t)$, the reader can easily verify that all the conditions of [23] are satisfied, as $\|\mathcal{Y}(X,V)\|_{\mathscr{S}^2}$ and $\|\mathcal{M}(X,V)\|_{\mathscr{S}^2}$ are finite by Lemma 4.2. On the other hand, it is easy to verify that $\widetilde{V}\in\mathscr{S}^2([0,T],\mathbb{R}^d)$, as for $f_0 := f(\cdot,0,0,0)$,

$$
\begin{aligned}
&\|\widetilde{V}\|_{\mathscr{S}^2} \leq \sqrt{T}\Big(\|f_0\|_{\mathscr{H}^2} + \|f\big(\cdot,\widetilde{X}_\cdot,\mathcal{Y}(X,V)_\cdot,\mathcal{L}^3(\mathcal{M}(X,V))_\cdot\big)-f_0\|_{\mathscr{H}^2}\Big)\\
&\leq \sqrt{T}\Big(\|f_0\|_{\mathscr{H}^2} + C(\sqrt{T}\vee 1)\big(\|\widetilde{X}\|_{\mathscr{S}^2} + \|\mathcal{Y}(X,V)\|_{\mathscr{S}^2} + \|\mathcal{L}^3(\mathcal{M}(X,V))\|_{\mathscr{O}^2}\big)\Big)\\
&\leq \sqrt{T}\Big(\|f_0\|_{\mathscr{H}^2} + C(\sqrt{T}\vee 1)\big(\|\widetilde{X}\|_{\mathscr{S}^2} + \|\mathcal{Y}(X,V)\|_{\mathscr{S}^2} + K\|\mathcal{M}(X,V)\|_{\mathscr{S}^2}\big)\Big) < \infty.
\end{aligned}
$$

Since the pair (X,V) is a solution of (3.3) if and only if it is a fixed point of $\mathbb{L}$, it suffices to prove that $\mathbb{L}$ is a contraction on $\mathscr{S}^2_X\times\mathscr{S}^2_V$ for small enough $T>0$. Let $\mathbb{L}(X^1,V^1)=(\widetilde{X}^1,\widetilde{V}^1)$, $\mathbb{L}(X^2,V^2)=(\widetilde{X}^2,\widetilde{V}^2)$, and assume without loss of generality that $T\leq 1$. By Itô's formula, we can compute that

$$
\begin{aligned}
d|\widetilde{X}^1_t-\widetilde{X}^2_t|^2 &= 2(\widetilde{X}^1_t-\widetilde{X}^2_t)^{\mathrm{T}} d(\widetilde{X}^1-\widetilde{X}^2)_t + d\langle \widetilde{X}^1-\widetilde{X}^2\rangle_t\\
&= 2\big(\widetilde{X}^1_t-\widetilde{X}^2_t\big)^{\mathrm{T}}\Big(\mu\big(t,\widetilde{X}^1_t,\mathcal{Y}(X^1,V^1)_t,\mathcal{L}^1(\mathcal{M}(X^1,V^1))_t\big)\\
&\qquad\qquad - \mu\big(t,\widetilde{X}^2_t,\mathcal{Y}(X^2,V^2)_t,\mathcal{L}^1(\mathcal{M}(X^2,V^2))_t\big)\Big)dt\\
&\quad + 2\big(\widetilde{X}^1_t-\widetilde{X}^2_t\big)^{\mathrm{T}}\Big(\sigma\big(t,\widetilde{X}^1_t,\mathcal{Y}(X^1,V^1)_t,\mathcal{L}^2(\mathcal{M}(X^1,V^1))_t\big)\\
&\qquad\qquad - \sigma\big(t,\widetilde{X}^2_t,\mathcal{Y}(X^2,V^2)_t,\mathcal{L}^2(\mathcal{M}(X^2,V^2))_t\big)\Big)dW_t\\
&\quad + \big|\sigma\big(t,\widetilde{X}^1_t,\mathcal{Y}(X^1,V^1)_t,\mathcal{L}^2(\mathcal{M}(X^1,V^1))_t\big)\\
&\qquad - \sigma\big(t,\widetilde{X}^2_t,\mathcal{Y}(X^2,V^2)_t,\mathcal{L}^2(\mathcal{M}(X^2,V^2))_t\big)\big|^2 dt
\end{aligned}
$$

for any $t\geq 0$. Thus, by applying Assumption (A1.2), (A1.4) and classical inequalities we obtain that, for a constant $\theta_1=\theta_1(C)$ depending only on C,

$$
\begin{aligned}
&\|\widetilde{X}^1-\widetilde{X}^2\|^2_{\mathscr{S}^2} = E\Big[\sup_{t\in[0,T]}|\widetilde{X}^1_t-\widetilde{X}^2_t|^2\Big]\\
&\leq \theta_1\Bigg(E\Big[\int_0^T|\widetilde{X}^1_s-\widetilde{X}^2_s|^2ds\Big] + E\Big[\int_0^T|\widetilde{X}^1_s-\widetilde{X}^2_s||\mathcal{Y}(X^1,V^1)_s-\mathcal{Y}(X^2,V^2)_s|ds\Big]\\
&\quad + E\Big[\int_0^T|\widetilde{X}^1_s-\widetilde{X}^2_s||\mathcal{L}^1(\mathcal{M}(X^1,V^1))_s-\mathcal{L}^1(\mathcal{M}(X^2,V^2))_s|ds\Big]\\
&\quad + E\Big[\int_0^T|\mathcal{Y}(X^1,V^1)_s-\mathcal{Y}(X^2,V^2)_s|^2ds\Big]\\
&\quad + E\Big[\int_0^T|\mathcal{L}^2(\mathcal{M}(X^1,V^1))_s-\mathcal{L}^2(\mathcal{M}(X^2,V^2))_s|^2ds\Big]\\
&\quad + E\Big[\Big(\int_0^T|\widetilde{X}^1_s-\widetilde{X}^2_s|^2\big(|\widetilde{X}^1_s-\widetilde{X}^2_s|^2 + |\mathcal{Y}(X^1,V^1)_s-\mathcal{Y}(X^2,V^2)_s|^2\\
&\qquad\qquad + |\mathcal{L}^2(\mathcal{M}(X^1,V^1))_s-\mathcal{L}^2(\mathcal{M}(X^2,V^2))_s|^2\big)ds\Big)^{1/2}\Big]\Bigg)\\
&=: \theta_1\Big(I_1+I_2+I_3+I_4+I_5+I_6\Big). \qquad (4.4)
\end{aligned}
$$

The next step consists in deriving estimates for all the terms on the right hand side of this inequality. I_1 can simply be estimated by

$$I_1 \le T\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2}. \tag{4.5}$$

For I_2, we obtain by the Cauchy-Schwarz inequality that

$$I_2 \le \frac{T}{2}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|\mathcal{Y}(X^1,V^1) - \mathcal{Y}(X^2,V^2)\|^2_{\mathscr{S}^2}\Big).$$

Because of assumption (A1.4) on ϕ, we can apply Lemma 4.2 and get that

$$I_2 \le \theta_2 T\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1 - X^2\|^2_{\mathscr{S}^2}\Big), \tag{4.6}$$

for some constant $\theta_2 = \theta_2(C)$ depending on C. To estimate I_3, one can apply again the Cauchy-Schwarz inequality to get that

$$\begin{aligned} I_3 &\le \frac{T}{2\varepsilon}\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \frac{\varepsilon}{2}E\Big[\int_0^T |\mathcal{L}^1(\mathcal{M}(X^1,V^1))_s - \mathcal{L}^1(\mathcal{M}(X^2,V^2))_s|^2 ds\Big] \\ &\le \frac{\sqrt{T}}{2}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|\mathcal{L}^1(\mathcal{M}(X^1,V^1)) - \mathcal{L}^1(\mathcal{M}(X^2,V^2))\|^2_{\mathscr{O}^2}\Big), \end{aligned}$$

where the last inequality is obtained by taking $\varepsilon = \sqrt{T}$ if $\mathscr{O}^2 = \mathscr{H}^2$ and $\varepsilon = 1$ if $\mathscr{O}^2 = \mathscr{S}^2$ (remember that $T \le 1$). On the other hand, by Assumption **(L1)** we know that

$$\|\mathcal{L}^1(\mathcal{M}(X^1,V^1)) - \mathcal{L}^1(\mathcal{M}(X^2,V^2))\|^2_{\mathscr{O}^2} \le K^2\|\mathcal{M}(X^1,V^1) - \mathcal{M}(X^2,V^2)\|^2_{\mathscr{S}^2},$$

and by applying Lemma 4.2,

$$I_3 \le \theta_3\sqrt{T}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1 - X^2\|^2_{\mathscr{S}^2}\Big), \tag{4.7}$$

for a constant $\theta_3 = \theta_3(C,K)$ depending only on C and K. I_4 is estimated by using the same argument as for (4.6), obtaining that

$$I_4 \le \theta_4 T\Big(\|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1 - X^2\|^2_{\mathscr{S}^2}\Big), \tag{4.8}$$

for some constant $\theta_4 = \theta_4(C)$. For I_5, we have that

$$\begin{aligned} I_5 &\le T\|\mathcal{L}^2(\mathcal{M}(X^1,V^1)) - \mathcal{L}^2(\mathcal{M}(X^2,V^2))\|^2_{\mathscr{S}^2} \\ &\le K^2 T\|\mathcal{M}(X^1,V^1) - \mathcal{M}(X^2,V^2)\|^2_{\mathscr{S}^2}, \end{aligned}$$

and hence, by Lemma 4.2,

$$I_5 \le \theta_5 T\Big(\|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1 - X^2\|^2_{\mathscr{S}^2}\Big), \tag{4.9}$$

for $\theta_5 = \theta_5(K)$. It only remains to estimate the last term I_6: for notational simplicity, we introduce the process $A^i := \Big(\mathcal{Y}(X^i,V^i), \mathcal{L}^2(\mathcal{M}(X^i,V^i))\Big)$ for $i = 1,2$. By applying the Cauchy-Schwarz inequality, we can verify that

$$\begin{aligned} I_6 &= E\Big[\Big(\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s|^2\big(|\widetilde{X}^1_s - \widetilde{X}^2_s|^2 + |A^1_s - A^2_s|^2\big)ds\Big)^{1/2}\Big] \\ &\le E\Big[\Big(\sup_{t\in[0,T]}|\widetilde{X}^1_t - \widetilde{X}^2_t|^2\Big)^{1/2}\Big(\int_0^T \big(|\widetilde{X}^1_s - \widetilde{X}^2_s|^2 + |A^1_s - A^2_s|^2\big)ds\Big)^{1/2}\Big] \\ &\le \frac{\sqrt{T}}{2}E\Big[\sup_{t\in[0,T]}|\widetilde{X}^1_t - \widetilde{X}^2_t|^2\Big] + \frac{1}{2\sqrt{T}}E\Big[\int_0^T \big(|\widetilde{X}^1_s - \widetilde{X}^2_s|^2 + |A^1_s - A^2_s|^2\big)ds\Big], \end{aligned}$$

and it is not difficult to check that

$$E\Big[\int_0^T (|\widetilde{X}^1_s - \widetilde{X}^2_s|^2 + |A^1_s - A^2_s|^2)ds\Big] \le T\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|A^1 - A^2\|^2_{\mathscr{S}^2}\Big),$$

which finally leads us to

$$I_6 \le \theta_6\sqrt{T}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1 - X^2\|^2_{\mathscr{S}^2}\Big), \tag{4.10}$$

for some constant $\theta_6 = \theta_6(C)$. Therefore, we can plug the estimates (4.5)–(4.10) back into (4.4), obtaining the inequality

$$\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} \le \theta_7\sqrt{T}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1 - X^2\|^2_{\mathscr{S}^2}\Big), \tag{4.11}$$

for some constant $\theta_7 = \theta_7(C, K)$. On the other hand, thanks to the explicit nature of the functional differential equation (4.3), we can estimate $\|\widetilde{V}^1 - \widetilde{V}^2\|_{\mathscr{S}^2}$ as follows:

$$\begin{aligned}
\|\widetilde{V}^1 - \widetilde{V}^2\|^2_{\mathscr{S}^2} \le E\Big[\Big(\int_0^T \Big|&f\big(s, \widetilde{X}^1_s, \mathcal{Y}(X^1, V^1)_s, \mathcal{L}^3(\mathcal{M}(X^1, V^1))_s\big)\\
&- f\big(s, \widetilde{X}^2_s, \mathcal{Y}(X^2, V^2)_s, \mathcal{L}^3(\mathcal{M}(X^2, V^2))_s\big)\Big|ds\Big)^2\Big]\\
\le TE\Big[\int_0^T \Big|&f\big(s, \widetilde{X}^1_s, \mathcal{Y}(X^1, V^1)_s, \mathcal{L}^3(\mathcal{M}(X^1, V^1))_s\big)\\
&- f\big(s, \widetilde{X}^2_s, \mathcal{Y}(X^2, V^2)_s, \mathcal{L}^3(\mathcal{M}(X^2, V^2))_s\big)\Big|^2 ds\Big]\\
\le 3C^2T\Big(&E\Big[\int_0^T |\widetilde{X}^1_s - \widetilde{X}^2_s|^2 ds\Big] + E\Big[\int_0^T |\mathcal{Y}(X^1, V^1)_s - \mathcal{Y}(X^2, V^2)_s|^2 ds\Big]\\
&+ E\Big[\int_0^T |\mathcal{L}^3(\mathcal{M}(X^1, V^1))_s - \mathcal{L}^3(\mathcal{M}(X^2, V^2))_s|^2 ds\Big]\Big)\\
\le 3C^2T\Big(&\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|\mathcal{Y}(X^1, V^1) - \mathcal{Y}(X^2, V^2)\|^2_{\mathscr{S}^2}\\
&+ \|\mathcal{L}^3(\mathcal{M}(X^1, V^1)) - \mathcal{L}^3(\mathcal{M}(X^2, V^2))\|^2_{\mathscr{O}^2}\Big),
\end{aligned}$$

and by using the same arguments as for the estimates (4.5)–(4.10) this yields that, for a constant $\theta_8 = \theta_8(C, K)$,

$$\|\widetilde{V}^1 - \widetilde{V}^2\|^2_{\mathscr{S}^2} \le \theta_8\sqrt{T}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1_t - X^2_t\|^2_{\mathscr{S}^2}\Big). \tag{4.12}$$

Hence, we can sum the inequalities (4.11) and (4.12), obtaining a constant $\theta_9 = \theta_9(C, K)$ depending only on C and K such that

$$\begin{aligned}
&\|(\widetilde{X}^1, \widetilde{V}^1) - (\widetilde{X}^2, \widetilde{V}^2)\|^2_{\mathscr{S}^2_X \times \mathscr{S}^2_V} \le \theta_9\sqrt{T}\Big(\|\widetilde{X}^1 - \widetilde{X}^2\|^2_{\mathscr{S}^2} + \|V^1 - V^2\|^2_{\mathscr{S}^2} + \|X^1 - X^2\|^2_{\mathscr{S}^2}\Big)\\
&\le \theta_9\sqrt{T}\Big(\|(\widetilde{X}^1, \widetilde{V}^1) - (\widetilde{X}^2, \widetilde{V}^2)\|^2_{\mathscr{S}^2_X \times \mathscr{S}^2_V} + \|(X^1, V^1) - (X^2, V^2)\|^2_{\mathscr{S}^2_X \times \mathscr{S}^2_V}\Big),
\end{aligned}$$

which implies that, for $T > 0$ such that $\theta_9\sqrt{T} < 1/2$,

$$\|(\widetilde{X}^1, \widetilde{V}^1) - (\widetilde{X}^2, \widetilde{V}^2)\|^2_{\mathscr{S}^2_X \times \mathscr{S}^2_V} \le \underbrace{\Big(\frac{1}{\theta_9\sqrt{T}} - 1\Big)^{-1}}_{<1} \|(X^1, V^1) - (X^2, V^2)\|^2_{\mathscr{S}^2_X \times \mathscr{S}^2_V}.$$

Therefore, $\mathbb{L}$ is a contraction if $T < \ell(C,K) := \frac{1}{4\theta_9^2(C,K)} \wedge 1$, and thus admits a unique fixed point (X,V). ☐

Remark 4.6. Under the general assumption **(A1)**, it is not possible to extend Theorem 4.5 to the case where $\mathcal{L}^2$ satisfies **(L1)** with respect to $\mathscr{O}^2([0,T],\mathbb{R}^{p_2}) = \mathscr{H}^2([0,T],\mathbb{R}^{p_2})$, as shown by the following counterexample borrowed from the theory of FBSDEs. Assume we have an augmented Brownian filtration, and consider the functional differential system

$$\begin{cases} dX_t = \mathcal{L}^2(\mathcal{M}^\phi(X,V))_t dW_t, \\ dV_t = 0, \\ X_0 = V_0 = 0, \end{cases}$$

where $\mathcal{L}^2$ is the operator given by Itô's representation, and $\phi(x) := x + W_T$. We thus obtain that $V \equiv 0$, and the first equation can be rewritten as

$$dX_t = d\mathcal{M}^\phi(X,0)_t, \quad X_0 = 0.$$

Assume it has an adapted solution X. We would then have $X_t = E[X_T + W_T|\mathcal{F}_t]$ for all $t \geq 0$, which would lead in particular, for $t = T$, to the contradiction $W_T = 0$.

Remark 4.7. While already quite general, Theorem 4.5 can be further extended in a number of different ways. We present some possible extensions which only require slight modifications of the proof presented above: the details are left to the reader.

(i) First of all, Theorem 4.5 can be extended to any initial time $\tau > 0$. More exactly, consider functional differential systems on $[\tau, T]$ of the form

$$\begin{cases} dX_t = \mu(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^1(\mathcal{M}(X,V))_t)dt + \sigma(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^2(\mathcal{M}(X,V))_t)dW_t, \\ dV_t = f(t, X_t, \mathcal{Y}(X,V)_t, \mathcal{L}^3(\mathcal{M}(X,V))_t)dt, \\ X_\tau = \eta, \quad V_\tau = \zeta, \end{cases}$$

where $\eta, \zeta \in L^2(\mathcal{F}_\tau)$. Then it is possible to prove that, under the same conditions of Theorem 4.5 and provided that $T - \tau < \ell$, there is a unique solution in $\mathscr{S}^2([\tau,T],\mathbb{R}^n) \times \mathscr{S}^2([\tau,T],\mathbb{R}^d)$. This remark will be essential in the next section, where we extend the solvability to any time interval.

(ii) It is not difficult to see that Theorem 4.5 also holds for operators $\mathcal{L}^1$ and $\mathcal{L}^3$ of the form $\mathcal{L}^i = (\mathcal{L}^i_1, \cdots, \mathcal{L}^i_{k_i}, \mathcal{L}^i_{k_i+1}, \ldots, \mathcal{L}^i_{l_i})$, $i = 1,3$, where $\mathcal{L}^i_j$ satisfies Assumption **(L1)** with respect to $\mathscr{S}^2\big([0,T],\mathbb{R}^{p^i_j}\big)$ for $1 \leq j \leq k_i$, and with respect to $\mathscr{H}^2\big([0,T],\mathbb{R}^{p^i_j}\big)$ for $k_i + 1 \leq j \leq l_i$: indeed, it suffices to consider the space $\mathscr{O}^2\big([0,T],\mathbb{R}^{p^i}\big) := \prod_{j=1}^{k_i} \mathscr{S}^2\big([0,T],\mathbb{R}^{p^i_j}\big) \times \prod_{j=k_i+1}^{l_i} \mathscr{H}^2\big([0,T],\mathbb{R}^{p^i_j}\big)$ and take the norm given by

$$\|L\|^2_{\mathscr{O}^2([0,T],\mathbb{R}^{p^i})} := \sum_{j=1}^{k_i} \|L\|^2_{\mathscr{S}^2([0,T],\mathbb{R}^{p^i_j})} + \sum_{j=k_i+1}^{l_i} \|L\|^2_{\mathscr{H}^2([0,T],\mathbb{R}^{p^i_j})}.$$

(*iii*) The result of Theorem 4.5 can also be extended to other types of terminal condition. Indeed, assume that $\mathbb{D}([0,T],\mathbb{R}^n)$ denotes the space of all $\mathbb{R}^n$-valued càdlàg functions on $[0,T]$, and let $\Phi:\Omega\times\mathbb{D}([0,T],\mathbb{R}^n)\to\mathbb{R}^d$ satisfy the L^∞-Lipschitz condition

$$|\Phi(\mathbf{x})-\Phi(\mathbf{x}')|\le C\sup_{t\in[0,T]}|\mathbf{x}_t-\mathbf{x}'_t|\quad P\text{-a.s.}\quad\forall\,\mathbf{x},\mathbf{x}'\in\mathbb{D}([0,T],\mathbb{R}^n),$$

where $C>0$. Then, the operators

$$\overline{\mathcal{M}}(X,V)_t := E[\Phi(X)+V_T|\mathcal{F}_t],\quad \overline{\mathcal{Y}}(X,V)_t := \overline{\mathcal{M}}(X,V)_t - V_t,$$

satisfy estimates similar to those of Lemma 4.2. Therefore, the reader can easily verify that Theorem 4.5 remains valid if we substitute the operators $\mathcal{M}$ and $\mathcal{Y}$ in the system (3.3) by $\overline{\mathcal{M}}$ and $\overline{\mathcal{Y}}$. Moreover, the same conclusion remains true if Φ satisfies the L^1-Lipschitz condition

$$|\Phi(\mathbf{x})-\Phi(\mathbf{x}')|\le C\int_0^T|\mathbf{x}_t-\mathbf{x}'_t|dt\quad P\text{-a.s.}\quad\forall\,\mathbf{x},\mathbf{x}'\in\mathbb{D}([0,T],\mathbb{R}^n),$$

instead of the above L^∞-Lipschitz condition. Two typical examples are the functionals $\Phi^1(\mathbf{x})=\sup_{t\in[0,T]}|\mathbf{x}_t|$ and $\Phi^2(\mathbf{x})=\int_0^T\mathbf{x}_t dt$, which are related to lookback and Asian options.

(*iv*) Finally, it is possible to generalize the system (3.3) by substituting the equation for the component V by

$$dV_t = f(t,X_t,\mathcal{Y}(X,V)_t,\mathcal{L}^3(\mathcal{M}(X,V))_t)\alpha_V(dt),\quad V_0=0,$$

where, for some $\widetilde{T}>0$, α_V is a non-random, finite Borel measure on $[0,\widetilde{T}]$ such that $\alpha_V(\{0\})=0$. This is done by applying a change of integration order similar to the one discussed in (4.1).

We conclude this section by observing that the flexibility in the choice of the operators $\mathcal{L}^i$ opens the door to probabilistic interpretations for many classes of integro-partial differential equations, similarly to the well known non-linear Feynman-Kac formula for BSDEs. We do not attempt to investigate this problem into more detail here, leaving it for future research.

5 Solvability on arbitrary intervals

After proving in Theorem 4.5 the local solvability of the fully coupled system (3.3), the next step consists in showing the existence and uniqueness of solutions on arbitrarily large time intervals. In the case of simple, non-coupled Lipschitz functional differential equations, Liang et al. [15] showed that the existence and uniqueness of solutions can be obtained by imposing additional conditions on the operator $\mathcal{L}$, while leaving the assumptions on the driver f and the terminal condition ξ unchanged.

However, the situation is quite different for coupled systems: as it is now well known in the theory of classical FBSDEs, an extension of Theorem 4.5 to arbitrary intervals is possible only if we impose additional assumptions on the coefficients μ, σ, f, ϕ of the fully coupled system (3.3), since the solution could explode for large time horizons (without going into more detail, we refer the reader to classical counterexamples for Markovian FBSDEs and related PDEs which can be found for instance in [18]). Several techniques have been adopted for classical FBSDEs to overcome this difficulty, as we mentioned in the Introduction. However, we prefer to apply another approach which in our opinion is the most natural in the case of functional differential systems.

First of all, we briefly discuss the intuition. Similarly to [15], the first step consists in dividing the interval $[0, T]$ into a finite number of subintervals $I_j := [T_{j-1}, T_j]$, $0 = T_0 < \cdots < T_N = T$: then, we solve the system separately on any subinterval, starting from the last one and going backward. There is, however, an important additional difficulty with respect to [15], which clarifies why additional assumptions on the coefficients are needed. For such an approach to work, we need in fact that the length of the subintervals I_j, on which the system has to be solvable, can be bounded by below by a constant independent of j. We have seen that such a length depends on the Lipschitz constants C and K of the system, and we can observe that the only potential complication can arise from the terminal condition on each I_j: indeed, while the other coefficients μ, σ, f and $\mathcal{L}^i$ remain the same on all intervals I_j, the terminal condition of the subsystem on I_j is given by $\Xi^j = \mathcal{Y}(X^{j+1}, V^{j+1})_{T_j}$, where (X^{j+1}, V^{j+1}) is the local solution on I_{j+1}. To obtain the desired lower bound for the length for all I_j, it is therefore sufficient that $\Xi^j = \theta_j(X^{j+1}_{T_j})$, where $\theta_j(\omega, x)$ is for each j Lipschitz continuous in x with respect to some constant C *independent of* j.

Due to the strong coupling between the two functional differential equations, this uniform Lipschitz continuity can be obtained only by studying the interplay between the two components X and V. However, this is very difficult without a concrete expression for the operators $\mathcal{L}^i$, and is especially true in the case of non-local operators (namely, the fact that $\mathcal{L}(M)$ could still depend on the whole path of $(M_t)_{t\in[0,T]}$ may easily cause the explosion of the solution for large time intervals). These observations lead us to the conclusion that the solvability on arbitrary intervals of the fully coupled system (3.3) has to be treated on a case-by-case basis, by developing tailor-made techniques for each choice of the operators $\mathcal{L}^i$.

As an example, we study in the following the case where the system (3.3) is associated to a classical Brownian FBSDE. Hence, we will assume for the rest of the article that the filtration is generated by an m-dimensional Brownian motion $(W_t)_{t\in[0,T]}$ on $(\Omega, \mathcal{F}, P)$, and that $\mathcal{L}^1 = \mathcal{L}^2 = \mathcal{L}^3 = \mathcal{L}$, where $\mathcal{L}$ is the operator given by Itô's representation theorem. For notational simplicity, we will write $\mathcal{Z}(X, V) = \mathcal{L}(\mathcal{M}(X, V))$.

In the literature of classical FBSDEs, it is critical to distinguish between the cases where the coefficients μ, σ, f and ϕ are purely deterministic (i.e. they do not depend on ω) or not. This is due to the fact that, in the purely deterministic case, one can exploit the well known connection between FBSDEs and parabolic PDEs. In the framework of functional differential equations, such an approach has been adopted by Liang et al. [14] in a very special case, where they essentially rely on the Lipschitz continuity of the solution of the corresponding parabolic PDEs (see for example Delarue [4, 5]).

We are, however, more interested in the case where the coefficients μ, σ, f and ϕ are allowed to be random. In the literature, there are several articles dedicated to the random case: in particular, Zhang [25, 26] proved the solvability of classical FBSDEs by deriving uniform Lipschitz estimates with respect to the initial condition of the component X. In the following, after reformulating in our setting Zhang's results on uniform Lipschitz continuity, we briefly illustrate how we can use these results together with Theorem 4.5 to construct a solution to our coupled functional differential system. *For the rest of this section, we will assume that $n = 1$, i.e. the component X is 1-dimensional, and that σ does not depend on z_2, i.e. $\sigma = \sigma(t, x, y)$.* Moreover, we assume that the coefficients of (3.3) satisfy the following condition:

Assumption (A2): *The functions $\mu : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}$, $\sigma : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \to \mathbb{R}^m$, $f : \Omega \times [0,T] \times \mathbb{R} \times \mathbb{R}^d \times \mathbb{R}^{d\times m} \to \mathbb{R}^d$ and $\phi : \Omega \times \mathbb{R} \to \mathbb{R}^d$ satisfy Assumption* **(A2)** *if they satisfy Assumption* **(A1)** *with* (A1.2) *replaced by* (A1.2'), *and:*

(A2.1) *ϕ is uniformly Lipschitz in x with constant C'.*

(A2.2) *There is a constant $\gamma > 0$ such that*

$$\Lambda_t^1(y) \le -\gamma|\Lambda_t^2(y)| \quad \forall\, y \in \{y' \in \mathbb{R}^d \,|\, |y'| = 1\}, \text{ where}$$

$$\begin{aligned}\Lambda_t^1(y) &:= \sum_{i=1}^d y_i\Big(\mathrm{Tr}\left(\partial_z f^i (\partial_z \mu)^{\mathrm{T}}\right) - y^{\mathrm{T}}\partial_z\mu(\partial_z f^i)^{\mathrm{T}} y + y^{\mathrm{T}}\partial_y\sigma(\partial_z f^i)^{\mathrm{T}} y\Big)\\ &\qquad + \partial_x\sigma(\partial_z\mu)^{\mathrm{T}} y + (\partial_y\mu)^{\mathrm{T}} y,\\ \Lambda_t^2(y) &:= |\partial_z\mu|^2 - |(\partial_z\mu)^{\mathrm{T}} y|^2 + 2y^{\mathrm{T}}\partial_z\mu(\partial_y\sigma)^{\mathrm{T}} y,\end{aligned}$$

and where we assumed that all corresponding derivatives exist.

The purpose of Assumption (A2.1) is to make a distinction between the Lipschitz constant of ϕ and those of the other coefficients, since such a distinction is needed hereinafter. Observe moreover that we require stronger regularity conditions than those of Delarue [4,5], since the coefficients of the system (3.3) are random (however, σ is allowed to be degenerate). Under these conditions, by deriving some clever estimates for linear FBSDEs, Zhang obtained the following result:

Lemma 5.1 (Zhang [26])**.** *Assume that the coefficients μ, σ, f and ϕ satisfy Assumption* **(A2)** *with $\gamma = \frac{1}{C}$, and let ℓ denote the constant in Theorem 4.5. Let $T < \ell$, and for $x^i \in \mathbb{R}$, $i = 1, 2$, let (X^i, V^i) denote the solution of*

$$\begin{cases} dX_t^i = \mu(t, X_t^i, \mathcal{Y}(X^i, V^i)_t, \mathcal{Z}(X^i, V^i)_t)dt + \sigma(t, X_t^i, \mathcal{Y}(X^i, V^i)_t)dW_t,\\ dV_t^i = f(t, X_t^i, \mathcal{Y}(X^i, V^i)_t, \mathcal{Z}(X^i, V^i)_t)dt,\\ X_0^i = x^i, \quad V_0^i = 0.\end{cases}$$

Then, there is a constant ϱ_C, depending only on C, such that

$$\big|\mathcal{Y}(X^1, V^1)_0 - \mathcal{Y}(X^2, V^2)_0\big| \le \overline{C}|x^1 - x^2|,$$

where $\overline{C} := \sqrt{\big(|C'|^2 + 1\big)e^{\varrho_C T} - 1} > 0.$

The proof can be found in [26]. As already mentioned, Lemma 5.1 in conjunction with Theorem 4.5 allows to construct a solution to the system (3.3) on arbitrarily large time intervals. More exactly:

Theorem 5.2. *Assume that the coefficients μ, σ, f and ϕ satisfy Assumption* **(A2)**. *Then, for any $T > 0$, the fully coupled system* (3.3) *has a unique solution (X, V) in $\mathscr{S}^2([0,T],\mathbb{R}^n) \times \mathscr{S}^2([0,T],\mathbb{R}^d)$.*

Proof. We assume without loss of generality that $\gamma = \frac{1}{C}$ in Assumption **(A2)**, by changing C or γ if necessary. Let $\overline{C}$ denote the constant in Lemma 5.1, and let $\ell = \ell(\overline{C})$ be the constant in Theorem 4.5. We consider a partition $(T_0, \cdots, T_N)$ of $[0,T]$ such that $0 = T_0 < \cdots < T_N = T$ and $0 < T_i - T_{i-1} < \ell$ for $i = 1, \cdots, N$, and set $I_i := [T_{i-1}, T_i]$. Moreover, to emphasize the dependence of the operators $\mathcal{Y}$, $\mathcal{M}$ on the terminal time and condition, we introduce a slightly different notation and write, for all $t \in [0, T_i]$,

$$\begin{aligned}
\breve{\mathcal{M}}^{T_i}(\xi, V)_t &= E[\xi + V_{T_i} | \mathcal{F}_t],\\
\breve{\mathcal{Y}}^{T_i}(\xi, V)_t &= E[\xi + V_{T_i} | \mathcal{F}_t] - V_t.
\end{aligned}$$

$\breve{\mathcal{Z}}^{T_i}(\xi, V)_t$ is then defined via the Itô representation of $(\breve{\mathcal{M}}^{T_i}(\xi, V)_t)_{t\in[0,T_i]}$.

The first step consists in constructing appropriate terminal conditions for all subintervals I_i. This is accomplished via a backward procedure. We set $\theta_N := \phi$, $C_N := C'$, and for all $x \in \mathbb{R}$, we consider the following system on I_N:

$$\begin{cases}
dX_t^{N,x} = \mu(t, X_t^{N,x}, \breve{\mathcal{Y}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_t, \breve{\mathcal{Z}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_t)dt\\
\qquad\qquad + \sigma(t, X_t^{N,x}, \breve{\mathcal{Y}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_t)dW_t,\\
dV_t^{N,x} = f(t, X_t^{N,x}, \breve{\mathcal{Y}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_t, \breve{\mathcal{Z}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_t)dt,\\
X_{T_{N-1}}^{N,x} = x, \quad V_{T_{N-1}}^{N,x} = 0.
\end{cases}$$

Since θ_N has Lipschitz constant $C_N \le \overline{C}$, the system has for all $x \in \mathbb{R}$ a unique solution $(X^{N,x}, V^{N,x})$ by Theorem 4.5 and Remark 4.7 (i). We can thus define θ_{N-1} by

$$\theta_{N-1}(x) := \breve{\mathcal{Y}}^{T_N}(\theta_N(X_{T_N}^{N,x}), V^{N,x})_{T_{N-1}}.$$

It is then easy to check that $\theta_{N-1}(x)$ is $\mathcal{F}_{T_{N-1}}$-measurable for all $x \in \mathbb{R}$. Moreover, by Lemma 5.1, θ_{N-1} is uniformly Lipschitz in x with constant

$$C_{N-1} := \sqrt{\big(|C_N|^2 + 1\big)e^{\varrho_C(T_N - T_{N-1})} - 1}.$$

Since $C_{N-1} \le \overline{C}$, we can iterate the same argument: for $i = N-1, \cdots, 2$, we consider for all $x \in \mathbb{R}$ the solution $(X^{i,x}, V^{i,x})$ on I_i of the system:

$$\begin{cases}
dX_t^{i,x} = \mu(t, X_t^{i,x}, \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_t, \breve{\mathcal{Z}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_t)dt\\
\qquad\qquad + \sigma(t, X_t^{i,x}, \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_t)dW_t,\\
dV_t^{i,x} = f(t, X_t^{i,x}, \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_t, \breve{\mathcal{Z}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_t)dt,\\
X_{T_{i-1}}^{i,x} = x, \quad V_{T_{i-1}}^{i,x} = 0.
\end{cases}$$

which exists by Theorem 4.5 and Remark 4.7 (i). We then define θ_{i-1} by

$$\theta_{i-1}(x) := \breve{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}^{i,x}), V^{i,x})_{T_{i-1}}.$$

$\theta_{i-1}(x)$ is thus $\mathcal{F}_{T_{i-1}}$-measurable for all $x \in \mathbb{R}$, and by Lemma 5.1, θ_{i-1} is uniformly Lipschitz in x with constant $C_{i-1} := \sqrt{\big(|C_i|^2 + 1\big)e^{\varrho_C(T_i - T_{i-1})} - 1}$. Moreover, we can easily verify by induction that

$$C_{i-1} = \sqrt{\big(|C_N|^2 + 1\big)e^{\varrho_C(T_N - T_{i-1})} - 1} \le \overline{C}.$$

Now that we have derived appropriate terminal conditions θ_i, we can construct the solution on the whole interval $[0,T]$ by a forward procedure. We set $X^0_{T_0} := x$, $V^0_{T_0} := 0$. Then, for $i = 1, \cdots, N$, we denote by (X^i, V^i) the solution on I_i of the system

$$\begin{cases} dX^i_t = \mu(t, X^i_t, \check{\mathcal{Y}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t, \check{\mathcal{Z}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t)dt \\ \qquad\quad + \sigma(t, X^i_t, \check{\mathcal{Y}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t)dW_t, \\ dV^i_t = f(t, X^i_t, \check{\mathcal{Y}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t, \check{\mathcal{Z}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t)dt, \\ X^i_{T_{i-1}} = X^{i-1}_{T_{i-1}}, \quad V^i_{T_{i-1}} = V^{i-1}_{T_{i-1}}. \end{cases}$$

which exists due to Theorem 4.5 and Remark 4.7 (i). We then set, for $t \in I_i$,

$$X_t := X^i_t, \quad V_t := V^i_t.$$

To prove that (X, V) solves (3.3) on $[0,T]$, it suffices to check that, for $t \in I_i$,

$$\check{\mathcal{Y}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t = \mathcal{Y}(\phi(X_{T_N}), V)_t, \quad \check{\mathcal{Z}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t = \mathcal{Z}(\phi(X_{T_N}), V)_t.$$

However, for $i = 1, \cdots, N-1$ and $t \in [0, T_i]$, we have that

$$\begin{aligned} \check{\mathcal{Y}}^{T_i}(\theta_i(X_{T_i}), V)_t &= \check{\mathcal{Y}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t \\ &= E\big[\mathcal{Y}^{T_{i+1}}\big(\theta_{i+1}(X^{i+1}_{T_{i+1}}), V^{i+1}\big)_{T_i} + V^i_{T_i}\big|\mathcal{F}_t\big] - V^i_t \\ &= E\big[\theta_{i+1}(X_{T_{i+1}}) + V_{T_{i+1}} - V_{T_i} + V_{T_i}\big|\mathcal{F}_t\big] - V_t \\ &= \check{\mathcal{Y}}^{T_{i+1}}(\theta_{i+1}(X_{T_{i+1}}), V)_t \end{aligned}$$

by the construction of θ_i. This gives by induction that $\check{\mathcal{Y}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t = \mathcal{Y}(\theta_N(X_{T_N}), V)_t = \mathcal{Y}(\phi(X_{T_N}), V)_t$ on I_i for $i = 1, \cdots, N$. On the other hand, this implies that $\check{\mathcal{M}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t = \mathcal{M}(\phi(X_{T_N}), V)_t$ on I_i for all i. In other words, $(\check{\mathcal{M}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t)_{t\in I_i}$ is the restriction to I_i of the martingale $(\mathcal{M}(\phi(X_{T_N}), V)_t)_{t\in[0,T]}$ and, due to the locality of the operator $\mathcal{Z}$, this gives us that $\check{\mathcal{Z}}^{T_i}(\theta_i(X^i_{T_i}), V^i)_t = \mathcal{Z}(\phi(X_{T_N}), V)_t$ on I_i.

This shows that (X, V) is a solution of (3.3) on $[0,T]$, and the proof is concluded by observing that the uniqueness is a consequence of the uniqueness of (X^i, V^i) on I_i. □

We conclude this article by briefly mentioning an extension of Zhang's results [25, 26] recently derived by Ma et al. [17]. Motivated by the connection between FBSDEs and PDEs in the deterministic case, the authors suggest that, for random coefficients, the solution can be extended to arbitrary intervals by relying on the existence of a random field θ (called decoupling field) such that $\mathcal{Y}(X, V)_t = \theta(t, X_t)$. It is then sufficient to show that θ is uniformly Lipschitz continuous. This can be obtained via the introduction of a backward stochastic Riccati equation of quadratic growth, under much weaker assumptions than Assumption **(A2)** (however, all processes have to be one-dimensional).

References

[1] F. Antonelli, Backward-forward stochastic differential equations, *Ann. Appl. Probab.*, **3**(3) (1993) 777–793.

[2] J.-M. Bismut, Conjugate convex functions in optimal stochastic control, *J. Math. Anal. Appl.*, **44** (1973) 384–404.

[3] M. Casserini and F. Delbaen, A note on quadratic BSDEs and associated functional differential equations, *ETH Zürich*, Forthcoming.

[4] F. Delarue, On the existence and uniqueness of solutions to FBSDEs in a non-degenerate case, *Stochastic Process. Appl.*, **99**(2) (2002) 209–286.

[5] F. Delarue, Estimates of the solutions of a system of quasi-linear PDEs: A probabilistic scheme, In *Séminaire de Probabilités XXXVII*, volume 1832 of *Lecture Notes in Math.*, Springer, Berlin, (2003) 290–332.

[6] Ł. Delong, Applications of time-delayed backward stochastic differential equations to pricing, hedging and portfolio management, *arXiv:1005.4417v3*, (2011).

[7] Ł. Delong and P. Imkeller, Backward stochastic differential equations with time delayed generators—results and counterexamples, *Ann. Appl. Probab.*, **20**(4) (2010) 1512–1536.

[8] G. Dos Reis, A. Réveillac, and J. Zhang, FBSDE with time delayed generators: L^p-solutions, differentiability, representation formulas and path regularity, *Stochastic Process. Appl.*, **121**(9) (2011) 2114–2150.

[9] N. El Karoui, S. Hamadène, and A. Matoussi, *Indifference pricing: theory and applications*, chapter BSDEs and applications, Princeton University Press, (2009) 267–320.

[10] N. El Karoui and L. Mazliak, editors: *Backward stochastic differential equations*, volume 364 of *Pitman Research Notes in Mathematics Series*, (Longman, Harlow, 1997).

[11] N. El Karoui, S. Peng, and M. C. Quenez, Backward stochastic differential equations in finance, *Mathematical Finance*, **7**(1) (1997) 1–71.

[12] U. Horst, Y. Hu, P. Imkeller, A. Réveillac, and J. Zhang, Forward-backward systems for expected utility maximization, *Stochastic Processes and their Applications*, Forthcoming.

[13] Y. Hu and S. Peng, Solution of forward-backward stochastic differential equations, *Probab. Theory Related Fields*, **103**(2) (1995) 273–283.

[14] G. Liang, T. Lyons, and Z. Qian: A functional approach to FBSDEs and its application in optimal portfolios, *arXiv:1011.4499v1*, (2010).

[15] G. Liang, T. Lyons, and Z. Qian, Backward stochastic dynamics on a filtered probability space, *The Annals of Probability*, **39**(4) (2011) 1422–1448.

[16] J. Ma, P. Protter, and J. M. Yong, Solving forward-backward stochastic differential equations explicitly—a four step scheme, *Probab. Theory Related Fields*, **98**(3) (1994) 339–359.

[17] J. Ma, Z. Wu, D. Zhang, and J. Zhang, On wellposedness of forward-backward SDEs—a unified approach, *arXiv:1110.4658v1*, (2011).

[18] J. Ma and J. Yong, *Forward-backward stochastic differential equations and their applications*, volume 1702 of *Lecture Notes in Mathematics*, (Springer-Verlag, Berlin, 1999).

[19] É. Pardoux and S. G. Peng, Adapted solution of a backward stochastic differential equation, *Systems Control Lett.*, **14**(1) (1990) 55–61.

[20] É. Pardoux and S. G. Peng, Backward stochastic differential equations and quasilinear parabolic partial differential equations, In *Stochastic partial differential equations and their applications (Charlotte, NC, 1991)*, volume 176 of *Lecture Notes in Control and Inform. Sci.*, Springer, Berlin, (1992) 200–217.

[21] É. Pardoux and S. Tang, Forward-backward stochastic differential equations and quasilinear parabolic PDEs, *Probab. Theory Related Fields*, **114**(2) (1999) 123–150.

[22] S. Peng and Z. Wu, Fully coupled forward-backward stochastic differential equations and applications to optimal control, *SIAM J. Control Optim.*, **37**(3) (1999) 825–843.

[23] B. L. Rozovsky, A note on strong solutions of stochastic differential equations with random coefficients, In *Stochastic differential systems (Proc. IFIP-WG 7/1 Working Conf., Vilnius, 1978)*, volume 25 of *Lecture Notes in Control and Information Sci.*, Springer, Berlin, (1980) 287–296.

[24] J. Yong, Finding adapted solutions of forward-backward stochastic differential equations: method of continuation, *Probab. Theory Related Fields*, **107**(4) (1997) 537–572.

[25] J. Zhang, The wellposedness of FBSDEs, *Discrete Contin. Dyn. Syst. Ser. B*, **6**(4) (2006) 927–940.

[26] J. Zhang, The wellposedness of FBSDEs (II), *Preprint.* Available at `http://www-bcf.usc.edu/~jianfenz/Papers/fbsde2.pdf`, (2006).